European Journal of Operational Research 144 (2003) 39–55

EUROPEAN JOURNAL OF OPERATIONAL RESEARCH

www.elsevier.com/locate/dsw

Discrete Optimization

Optimizing the use of contingent labor when demand is uncertain

Edieal J. Pinker [a,*], Richard C. Larson [b]

[a] W.E. Simon Graduate School of Business, University of Rochester, Rochester, NY 14627, USA

[b] Operations Research Center, Massachusetts Institute of Technology, Cambridge, MA 02139, USA

Received 19 December 2000; accepted 26 October 2001

Abstract

We develop a new model for flexible workforce management in environments with uncertainty in the demand for labor. In particular we model a workforce comprised of regular workers who have fixed schedules and may work overtime, and contingent workers whose working hours are flexible as specified by a contract over a finite planning horizon. Our model can represent a variety of contracts, including: temporary workers, on-call workers with guaranteed minimum pay, and comp-time arrangements. We formulate the model as an optimization problem that determines the regular and contingent worker pool sizes that minimize expected labor and backlog costs. Embedded within this problem is a dynamic programming problem of making optimal operational staffing decisions with respect to the utilization of contingent and overtime resources. Numerical examples demonstrate the effect that the timing of information has on the benefits of flexibility. We also derive structural results that may be exploited to reduce the computational effort required to use the model.

Keywords: Manpower planning; Human resources; Dynamic programming

1. Introduction and literature review

Between 1990 and 2000 revenue in the temporary help industry grew from 17.0 to 63.6 billion dollars (not inflation adjusted), and annual employment increased from .99 million to 2.54 million (American Staffing Association, 2001). This rapid growth is but one sign of the dramatic changes the US labor market has been undergoing in the past decade. While the US labor market has traditionally been a leader in labor flexibility the same trends are appearing in Europe as well. For example, in the year 2000, 60% of the sales of Manpower, the leading staffing services provider, came from its European operations (Manpower Inc., 2000). Throughout the economy, especially in services, firms rely more and more upon external contract workers to fill key positions that were once considered the exclusive purview of full-time permanent workers

* Corresponding author. Tel.: +1-716-275-2131; fax: +1-716-273-1140.
E-mail address: pinker@simon.rochester.edu (E.J. Pinker).

0377-2217/03/$ - see front matter
PII: S0377-2217(01)00378-2

Nomenclature

Parameters

V	number of periods in planning horizon
G	guaranteed fraction of V periods contingent workers are paid for
c_{rw}, c_{cw}, c_{ot}	cost per shift of regular workers, contingent workers and overtime work, respectively
c_{b}	cost of a unit of backlog or unsatisfied demand per period
c_{B}	cost of a unit of backlog or unsatisfied demand in the final period
c_{f}	fixed cost per contingent worker per planning horizon
π_{rw}, π_{cw}, π_{ot}	productivity of regular workers, contingent workers, and overtime work in units of work per shift, respectively
$\mathrm{OT}_{\mathrm{cw}}$, $\mathrm{OT}_{\mathrm{rw}}$	parameters determining upper limit of overtime shifts contingent and regular workers may work in a period, respectively
$\mathrm{OT}(S)$	upper bound on the total overtime shifts in a period for staff $S = (N, M)$

Decision variables

N, M	number of regular and contingent workers staffed for the planning horizon
ω_t	the number of overtime shifts utilized in even stages t
u_t	the number of contingent worker shifts utilized in odd stages t

Random variables

n_t, $\bar{N}$	actual and expected number of regular workers present in stage t
x_t	work in system at start of stage t, the net productivity of staff already committed, in units of regular worker productivity in a shift
κ_t	unused portion of the total contingent worker guarantee at end of stage $t-1$
ϕ_t	the amount of new work arriving to the system in stage t for odd t, in units of regular worker productivity in a shift
θ_t	the amount of new work arriving to the system in stage t for even t, in units of regular worker productivity in a shift

(Applebaum, 1987; Clinton, 1997; Brotherton, 1995). At the same time a variety of flexible staffing arrangements have become more prevalent within firms. Some of these arrangements include part-time, comp-time [1] and on-call work (Christensen, 1995). The causes of these changes are quite varied and complex and, being recent phenomena, are still under investigation. In fact the US Bureau of Labor Statistics only began collecting data on contingent workers in 1995 (US Bureau of Labor Statistics, 1995). When the adoption of these different internal and external staffing arrangements is driven by a firm's uncertainty in its demand for labor we refer to them as contingent work arrangements, using the definition in Polivka (1989): "Any job in which an individual does not have an explicit or implicit contract for long-term employment or one in which the minimum hours worked can vary in a non-systematic manner".

As a result of the above changes in the labor market, managers now have many more choices in crafting a staffing strategy that flexibly matches labor supply to demand. In this paper we formulate and analyze an operational model of a firm's use of contingent labor that provides managers with a tool to determine how many full-time and contingent workers to hire. Furthermore we provide a framework for quantifying the effects of having more or less demand information when making short-term staffing adjustments. We be-

[1] Comp-time is the practice by which workers are compensated for overtime in one week by shorter hours in the following week.

lieve such a tool is important for two major reasons. First, there is much anecdotal evidence that firms use contingent workers and find that they do not reduce costs (Nollen and Axel, 1995). This suggests either that they do not know how to use contingent workers effectively, or that they are not able to accurately evaluate the benefits of doing so. Second, from the authors' experiences working with the United States postal service (USPS) and a financial services firm it is apparent to us that without the correct information systems in place to track labor demand, it is difficult to take advantage of the flexibility of contingent workers. Our numerical experiments reveal three interesting insights. First, we see that information and flexibility strengthen each other's cost reduction effects. Second, we see that information and capacity act as complements, in the economic sense, in that increasing the demand information available increases the staffing levels selected. Thirdly, we see that increased contingent labor flexibility does not always decrease regular worker staffing levels.

We formulate a model of contingent labor management based upon our observations of the back room operations of various services. We consider a workplace in which there is uncertainty in the demand for labor resources (on a daily or weekly basis), there is uncertainty in the supply of labor (absenteeism), work can be backlogged, and no inventory of processed work is allowed. We assume that the employer may draw upon three sources of labor: fixed, over-time, and contingent, to minimize costs of labor and backlogged work. This description could apply to many clerical operations (mail sorting, insurance claims processing etc.), delivery services, and repair services among others. The key characteristic is that work is created by a customer request. Explicit to our model is a representation of how demand information becomes available to the decision maker and a model of contingent labor contracts.

A single-period model of the use of contingent workers appears in Abraham (1986) assuming an unlimited supply of flexible labor to a firm and no particular structure to the relationship. Berman and Larson (1994) model the use of contingent workers in an environment where demand is realized each day before decisions on the use of contingent workers must be made, there is no bound on overtime each day and no backlogging of unfinished work. A hierarchical model for planning the use of long-term, medium-term, and short-term flexibility is presented in Wild and Schneeweiss (1993). That paper models the use of short-timing (cutting back full-time hours), temporary workers, overtime, and crosstrained workers in an environment with uncertain demand for labor and absenteeism. It assumes that there is no backlogging or inventories and therefore daily staffing decisions are independent of one another. The model demonstrates the importance of the interplay between long-term planning decisions and the availability of resources for short-term staffing decisions. The model does not, however, represent different types of contingent worker contracts and it does not represent the way demand information timing affects their utilization. Our paper addresses both of these issues.

There has been recent interest in modeling contracts between suppliers and buyers within the inventory control and supply chain management literature (see Donohue, 2000; Anupindi and Bassok, 1998, 1997; Chen et al., 2001 for example). While these models have shed some light on some of the important issues of supplier–buyer relations, they are not structured to be directly applied to the workforce management setting. An exception is Milner and Pinker (2001) which considers labor supply contracts but not at the tactical decisionmaking level done here. This particular combination of overtime, processing capacity constraints, guaranteed work, and lack of inventories has not been addressed in the broader production literature.

In Sections 2 and 3 we, respectively, formulate and analyze our model. In Section 4 we present a set of illustrative numerical results. We conclude in Section 5. For the reader's convenience a glossary of notation used in the model is also given.

2. Model formulation

We study a firm that must set staffing levels for a planning horizon of V periods. The firm decides how many regular workers N to hire and how many contingent workers M to contract from a contingent labor

supplier. Regular workers are each guaranteed a wage (including benefits) of c_{rw} per period. The cost of the contingent labor is established by the contract. The contract specifies how many contingent workers M are contracted, a fixed payment of c_{f} per worker, a wage rate (including benefits) of c_{cw} per period worked by each worker and a guaranteed number of person-periods of work MGV, where G is $[0, 1]$.

In this section we formulate a mathematical model of the firm's staffing problem that includes the use of regular workers, contingent workers, and overtime and allows for uncertainty in demand for labor, backlogging of unfinished work, absenteeism, differing productivity rates across classes of workers, and different demand information scenarios. To properly assess the actual labor and backlog costs incurred by a particular solution, we model the operation of the workplace with optimal labor resource allocation decisions made on the period level. We model this level of activity with a dynamic programming model in which the manager has two decisions each period: how many contingent workers to use and how much overtime to use. Over the entire planning horizon these decisions determine the total expected labor and backlog costs for the system. In each period new work arrives to the system and is combined with the existing backlog. The manager then draws upon the three labor sources to process the work. The regular workers are automatically used, while contingent worker and overtime usage are both decided by the manager. Unprocessed work in a period is backlogged to the next period. Since the manager may have to decide how many contingent workers to call upon before deciding how much overtime to use the manager may have more information about the period's demand when making the latter decision.

2.1. Definitions and assumptions

Order of events. We define a time *period* to be an arbitrary, problem specific, unit of time. As discussed in Section 1, the model developed in this paper is geared toward periods of duration of the order of days or weeks, and within each period two distinct staffing decisions are made, how many contingent workers to use and how much overtime work to use. We assume the following sequence of events during each period:

- A first batch of work arrives and regular worker absenteeism is realized.
- The contingent worker usage decision is made.
- A second batch of work arrives.
- The overtime decision is made.

By "work arrival" we mean demand information rather than physical arrival of work. Some initial information about the quantity of new work for a period may arrive before the contingent worker decision is made, while additional information arrives after the decision is made. For example at the end of one work day (Monday) a manager may know how many regular workers have asked for the following day (Tuesday) off and have a partial forecast of how much work there will be the next day (Tuesday). At this point she can decide how many contingent workers to call in advance to come to work the next day (Tuesday). During Tuesday the manager will see the true extent of the workload and can decide to use overtime. The shorter the notice the manager has to provide contingent workers the more demand information available to her when she makes that decision. To reflect the fact that each of the staffing decisions may be made with differing amounts of demand information and absenteeism information we split each period i into a first or odd stage $t = 2i - 1$ and a second or even stage $t + 1 = 2i$, where $i = 1, \ldots, V$ with $t = 1, \ldots, 2V$.

For an odd stage t one decides how many contingent workers to use, $u_t = u_t(n_t, \phi_t)$, where n_t (the number of regular workers present) and ϕ_t (the first arrival of work in the period) are realized quantities. For an even stage t one decides how much overtime work to use, $\omega_t = \omega_t(\theta_t)$, where θ_t (the second arrival of work in the period) is the realized quantity.

Note that it is arbitrary to require the absenteeism of regular workers to be known before contingent worker decisions are made. The model could be formulated, with few changes, with the absenteeism in-

formation becoming available after the contingent worker decision is made (see Pinker, 1996). The alternative formulation leads to similar results.

Demand process. We assume that the ϕ_t's are independent, that all the θ_t's are independent as well, and that the ϕ_t's and θ_t's are independent of each other. To simplify notation in this paper we assume that the ϕ_t's are identically distributed as are the θ_t's, although the model could accommodate non-homogenous demand processes. The sum of the two random variables ϕ_{t-1} and θ_t defines the total exogenous demand for labor in the period spanning the stages $t-1$ and t, i.e., period $t/2$.

Absenteeism. We assume that the distribution of the number of regular workers present in a period is the same in all periods. We also assume that regular workers are paid for periods they are absent. This is similar to practice, in that workers are given an allowance of sick and personal days. We also assume that there is no absenteeism among contingent workers. This means that the contingent labor supplier is compelled to always provide up to M contingent workers by the contract.

Overtime. We consider any hours attributed to contingent workers after u_t has been reached, i.e., in even stages, as overtime and not counted toward the guarantee. For even stages t we decide ω_t, the number of overtime shifts utilized in stage t. Overtime each period is constrained such that

$$\omega_t \leqslant \mathrm{OT}_{\mathrm{rw}}\bar{N} + \mathrm{OT}_{\mathrm{cw}}M = \mathrm{OT}(S),$$

where $\bar{N}$ is the expected number of regular workers who are present in a period, and $\mathrm{OT}_{\mathrm{rw}}$ and $\mathrm{OT}_{\mathrm{cw}}$ are parameters, in units of shifts, that determine the maximum allowable overtime an individual worker of each type may perform in one period. Since the bound on overtime is expressed in terms of the expected number of regular workers present and not the actual number, there can be cases in which individuals exceed the overtime bound. Therefore, the overtime bound is an approximation of practice, made in the interest of tractability. We assume that contingent and regular workers both have the same productivity, π_{ot}, when performing overtime, also in the interest of tractability.

Cost parameters. We use c_{f} to denote the fixed cost for each contingent worker that is a member of the workforce for the planning period. It includes fixed component of compensation and fixed costs per worker for the firm, e.g., human resource departments costs, services available to all employees regardless of status, etc. Let c_{rw} be the per-shift cost of a regular worker that combines the benefits and salary. Let c_{cw} be the per-shift cost of a contingent worker that combines the prorated benefits and salary. Let c_{ot} be the cost per-worker per shift of overtime worked. Let c_{b} be the penalty per period incurred by the firm for every unit of work backlogged in the first $V-1$ periods, and let c_{B} be the corresponding penalty for the final period of the planning horizon. We allow for a different backlog penalty in the final period to provide flexibility to the user of this model.

2.2. Pool sizing problem formulation

We now formulate the pool sizing problem as

$$\begin{aligned}
\text{(P)} \quad \min_{S \geqslant 0} \quad & C(S) = c_{\mathrm{rw}}NV + c_{\mathrm{f}}M + c_{\mathrm{cw}}MGV + E_{\phi_1,n_1}\{f_1(0, MGV)\} \\
\text{s.t.} \quad & S = (N, M) \text{ integer} \quad \text{and} \quad B(S) \geqslant (0,0),
\end{aligned}$$

where $B(S)$ is a set of side constraints that may restrict the proportion of workers that may be of one type or the other. The cost of hiring N regular workers is $c_{\mathrm{rw}}NV$, and $c_{\mathrm{f}}M + c_{\mathrm{cw}}MGV$ is the up front cost of M contingent workers. The expression $E_{\phi_1,n_1}\{f_1(0, MGV)\}$ is the expected cost of making optimal tactical staffing decisions over the planning horizon given a staffing level $S=(N,M)$, starting with 0 units of work in the system, and MGV guaranteed shifts of contingent workers unused in stage 1. The state of the system in stage t is defined by x_t units of work (net productivity of committed staff) and κ_t guaranteed shifts of

contingent labor. In the following we use the notation $(y)^+$ and $[y]^+$ interchangeably to denote the positive part of y. To simplify the notation we define a new random variable $\phi'_t = \phi_t - \pi_{\rm rw} n_t$, which is new work arrival in odd stages net work done by the regular workers who are present. ϕ'_t can take on negative values. We can then define $f_t(x_t, \kappa_t)$ recursively as follows.

The cost-to-go function in stage $2V$ is:

$$f_{2V}(x_{2V}, \kappa_{2V}) = \min_{\omega_{2V}} \{c_{\rm ot}\omega_{2V} + c_{\rm B}(x_{2V} - \pi_{\rm ot}\omega_{2V})\},$$
$$\omega_{2V} \leqslant \min[x_{2V}/\pi_{\rm ot}, {\rm OT}(S)], \text{ where } {\rm OT}(S) = {\rm OT}_{\rm rw}\bar{N} + {\rm OT}_{\rm cw}M.$$

For even t:

$$f_t(x_t, \kappa_t) = \min_{\omega_t} \{c_{\rm ot}\omega_t + c_{\rm b}(x_t - \pi_{\rm ot}\omega_t) + E_{\phi'_{t+1}}[f_{t+1}(x_{t+1}, \kappa_{t+1})]\},$$
$$\omega_t \leqslant \min[x_t/\pi_{\rm ot}, {\rm OT}(S)], \text{ where } {\rm OT}(S) = {\rm OT}_{\rm rw}\bar{N} + {\rm OT}_{\rm cw}M.$$

In even stages t the cost-to-go is a combination of overtime charges $c_{\rm ot}$, backlog penalties $c_{\rm b}$ and expected future costs of operating the system.

For odd t: $$f_t(x_t, \kappa_t) = \min_{u_t} \left\{c_{\rm cw}[u_t - \kappa_t]^+ + E_{\theta_{t+1}}[f_{t+1}(x_{t+1}, \kappa_{t+1})]\right\}, \quad 0 \leqslant u_t \leqslant M.$$

In odd stages t the cost-to-go when u_t contingent workers are solicited is determined by the number of remaining guaranteed shifts plus the expected future costs of operating the system.

From even stage t to odd stage $t+1$ state transitions occur as follows:

$$x_{t+1} = x_t - \pi_{\rm ot}\omega_t + \phi'_{t+1} \quad \text{and} \quad \kappa_{t+1} = \kappa_t.$$

From odd stage t to even stage $t+1$ state transitions occur as follows:

$$x_{t+1} = [x_t - \pi_{\rm cw}u_t + \theta_{t+1}]^+,$$
$$\kappa_{t+1} = [\kappa_t - u_t]^+.$$

We have defined a two-dimensional state space (see Fig. 1) (x_t, κ_t), where κ_t is the unused portion of the total contingent worker guarantee, MGV, at the end of stage $t-1$, and x_t is the work in the system at the start of stage t, net the productivity of staff already committed for stage t. Note x_t may take on negative values for odd t. If x_t is positive it means that the backlogged work in the system exceeds the regular staff processing capacity. If x_t is negative it means that there are more regular workers present than needed for the backlog carried over from stage $t-1$ and therefore the staff can be assigned to work on the exogenous work arriving to the system in the second stage.

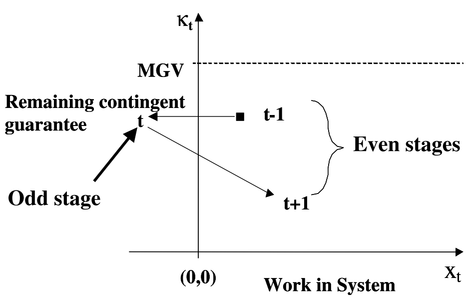

Fig. 1. State space for tactical problem dynamic program.

3. Analysis

Solving the staffing problem requires the determination of a cost minimizing staffing level $S = (N, M)$, where the evaluation of a particular solution S involves the optimal solution of a stochastic dynamic programming problem. In this section we present results that simplify the computational effort necessary for the dynamic optimization. The key results are summarized as follows:

- The cost-to-go functions of the dynamic optimization are convex in the state variables.
- Finding the optimal overtime and contingent worker utilization policies are convex minimization problems in each stage.
- There are easily computed bounds on the optimal u_t and w_t in each period.

The key to efficiently optimizing the model is to take advantage of the fact that the size of the dynamic program embedded within the model is directly proportional to the size of the contingent labor pool M and the guarantee G being evaluated. Therefore we prefer to evaluate the DP for smaller values of MG. The results we derive in this paper allow us to calculate bounds on the decision variables M, u_t, and ω_t using only calculations in which the κ dimension of the DP state space is eliminated.

3.1. Dynamic optimization

Proposition 1. *$f_t(x_t, \kappa_t)$ is convex, in (x_t, κ_t), for all t, and the choice of u_t and ω_t are convex minimization problems for all t.*

Proof. See Appendix A. □

3.1.1. Overtime policy

For even t we can rewrite the cost-to-go function as

$$f_t(x_t, \kappa_t) = L(x_t, \kappa_t) + \frac{c_{\mathrm{ot}}}{\pi_{\mathrm{ot}}} x_t,$$

where

$$L(x_t, \kappa_t) = \min_{[x_t - \mathrm{OT}(S)\pi_{\mathrm{ot}}]^+ \leqslant b_t \leqslant x_t} \{\Gamma_t(b_t, \kappa_t)\},$$

$$\Gamma_t(b_t, \kappa_t) = \left(c_{\mathrm{b}} - \frac{c_{\mathrm{ot}}}{\pi_{\mathrm{ot}}}\right) b_t + E_{\phi'_{t+1}}[f_{t+1}(b_t + \phi'_{t+1}, \kappa_{t+1})]$$

and

$$b_t = x_t - \pi_{\mathrm{ot}} \omega_t.$$

The change of variables means that selecting the optimal amount of overtime to use is equivalent to selecting an optimal amount of backlog, b_t, to accept or tolerate in the system. Let us denote the backlog that minimizes $\Gamma_t(b_t, \kappa_t)$ for a particular κ_t in period t to be $\beta_t(\kappa_t)$ which exists due to the convexity of $f_t(x_t, \kappa_t)$. That is, $\beta_t(\kappa_t)$ is the minimizer of $\Gamma_t(b_t, \kappa_t)$ without regard to the overtime usage constraint in period t or the work in the system x_t. Once we have determined $\beta_t(\kappa_t)$, the optimal overtime $\omega_t^*(x_t, \kappa_t)$ for a particular state is found by comparing $\beta_t(\kappa_t)$ with x_t and $[x_t - \pi_{\mathrm{ot}}\mathrm{OT}(S)]^+$ as follows:

Lemma 1. *Given $\beta_t(\kappa_t)$ and $x_t > 0$ the optimal overtime usage is*: $\omega_t^* = (x_t - b_t^*)/(\pi_{\text{ot}})$, *where*

$$b_t^*(x_t,\kappa_t) = \begin{cases} x_t & \text{if } \beta_t(\kappa_t) \geqslant x_t, \\ \beta_t(\kappa_t) & \text{if } [x_t - \pi_{\text{ot}}\text{OT}(S)]^+ \leqslant \beta_t(\kappa_t) < x_t, \\ [x_t - \pi_{\text{ot}}\text{OT}(S)]^+ & \text{if } \beta_t(\kappa_t) < [x_t - \pi_{\text{ot}}\text{OT}(S)]^+. \end{cases}$$

Proof. Follows from Proposition 1 and discussion above. □

When $c_{\text{b}} \geqslant c_{\text{ot}}/\pi_{\text{ot}}$, then $\beta_t(\kappa_t) = 0$ for all κ_t because the per period cost of a unit of backlog is greater than the overtime cost of eliminating it. Using the above characterization of the optimal overtime policy we can prove that the following bound on ω_t holds.

Proposition 2. *When* $c_{\text{b}} < c_{\text{ot}}/\pi_{\text{ot}}$,

$$\omega_t^* \leqslant (x_t - b_t^*(x_t, 0))/\pi_{\text{ot}}.$$

Proof. See Appendix A. □

3.1.2. Contingent worker policy

We note that the size of the state space of the model increases with MGV. In particular when $G = 0$ we have a one-dimensional problem that is solvable relatively quickly. We use the optimal contingent worker utilization policy from this special case to compute lower bounds on contingent worker utilization in general. We first state the following preliminary result regarding the optimal u_t.

Proposition 3. *When $\kappa_t = 0$ the optimal contingent worker usage in odd stages t, U_t, is characterized as follows*:

$$U_t(x_t) = \begin{cases} 0 & \text{if } x_t \leqslant X_t, \\ \frac{x_t - X_t}{\pi_{\text{cw}}} & \text{if } X_t + M\pi_{\text{cw}} \geqslant x_t > X_t, \\ M & \text{if } x_t > X_t + M\pi_{\text{cw}}, \end{cases}$$

where X_t is the minimizer of $(-c_{\text{cw}}/\pi_{\text{cw}})X_t + E_{\theta_{t+1}} f_{t+1}([X_t + \theta_{t+1}]^+, 0)$.

Proof. See Appendix A. □

The X_t defined in Proposition 3 can be computed using standard backwards recursion. If $\kappa_t > 0$, then contingent workers are cheaper to use than if $\kappa_t = 0$, and therefore, $u_t^*(x) \geqslant U_t(x)$ for all values of κ_t, i.e., the $U_t(x)$ are an easily computed lower bound on $u_t^*(x)$ in each odd stage.

3.2. Pool-sizing problem

In our experience with minimizing the objective function of problem (P), $C(S)$, we have found that it is convex and therefore a gradient search or binary search procedure always converges. However, we have not been able to prove this mathematically. Intuitively one would expect that the cost reduction of adding additional staff on the dynamic programming component of the objective function diminishes as S increases.

We also make the following observations without proof:

(O1) The optimal regular worker staff size N^* should be bounded above by the optimal N when no contingent workers are available, $N^*(M=0)$.

(O2) The following condition: $\bar{N} + \pi_{cw} M + \pi_{ot}\mathrm{OT}(S) \geqslant E[\phi_t + \theta_t]$ should be true to prevent instability.

(O3) It should be the case that $M^*(N) \leqslant M^*(N-1)$, where $M^*(N)$ is the optimal value of M given a fixed N.

The implication of (O3) is that by searching through the space of feasible S starting with large values of N and small values of M we can improve the lower bound on $M^*(N)$ with solutions to smaller problems.

4. Numerical results

In this section we conduct a set of numerical experiments in which we vary the amount of demand information available to the manager when he makes contingent worker utilization decisions and we vary the flexibility of the workers. We investigate how the firm's costs and the optimal staffing levels are affected by information and flexibility.

We assume that the firm contracts contingent labor from an agency that requires notification of a particular day's labor requirement by a fixed deadline. We assume the total amount of work that arrives each day is a Poisson random variable with mean 20. The agency's notification time splits the total amount of work into two batches which we assume are independent Poisson random variables with means that sum to 20. The larger the mean of Batch 1 the greater the information available to the manager when notifying the agency, in the sense that the variance in the total demand given the Batch 1 arrival is smaller. Similarly, the later the notification time the larger the mean of Batch 1. We consider five information scenarios with the following (Batch 1, Batch 2) means: $(20,0), (14,6), (10,10), (6,14)$, and $(0,20)$. We vary flexibility by adjusting the contingent worker guarantee G. For each information scenario we consider eight possible values for G, 0, .2, .3, .4, .5, .6, .7, and .8.

For the purposes of these experiments we model the regular worker absenteeism with a binomial distribution. For odd stages t, $n_t \sim \mathrm{Binomial}(N, p_N)$ is the number of regular workers present in stage t. We are assuming that each worker is present with a probability p_N, independently of all other workers.

In all the cases we assume the following parameters values: $V = 20$, $c_{rw} = 1$, $c_{cw} = 1.2$, $c_{ot} = 2.0$, $c_f = 0$, $c_b = 2.5$, $c_B = 2.5$, $\mathrm{OT}_{cw} = .25$, $\mathrm{OT}_{rw} = .25$, $\pi_{cw} = \pi_{ot} = 1.0$, $p_N = .95$.

In Fig. 2 we plot the firm's total cost under optimal staffing as a function of the contingent worker guarantee for the five different information scenarios. We can see that cost increases with decreased flexibility (as G increases) and costs decrease as more demand information is available earlier. We note that the

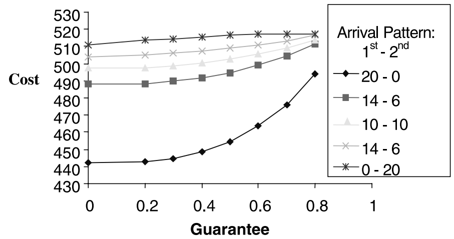

Fig. 2. Total firm costs for different information scenarios and flexibility levels.

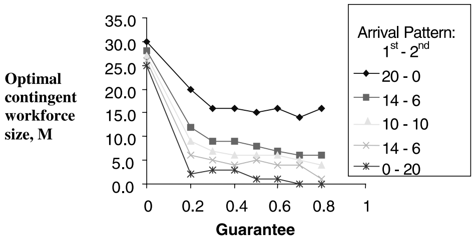

Fig. 3. Contingent labor staffing levels for different information scenarios and flexibility levels.

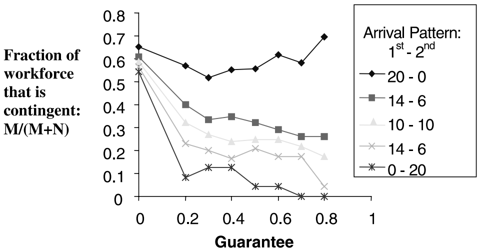

Fig. 4. Contingent fraction of workforce for different information scenarios and flexibility levels.

cost to the firm of staffing with only regular workers is approximately 517 and this places an upper bound on costs with contingent workers available. Interestingly we see that flexibility and information each enhance the cost reductions brought about by the other.

In Figs. 3 and 4 we plot the number of contingent workers contracted and the fraction of the total workforce that is contingent, respectively, as a function of the contingent worker guarantee for the five different information scenarios. Looking at these two figures together we can see that as information increases the firm contracts more contingent workers and they make up a greater fraction of the workforce for all levels of flexibility. This is interesting because it means that the firm reserves more of the flexible capacity as it has more access to demand information. In other words information and capacity are acting as complements. This result makes sense because additional information is making each unit of flexible capacity more effective. We also see in Fig. 3 that as flexibility increases (i.e., G decreases) the optimal contingent workforce size tends to increase. [2] We note that even when contingent worker decisions are made and no demand information is available, the 0–20 scenario, it is optimal to have some contingent workers to work down backlog.

In Fig. 5 we plot the number of regular workers staffed as a function of the contingent worker guarantee for the five information scenarios. We see that as flexibility of the contingent worker decreases (i.e., G increases) regular full time staffing may decrease. While it is true that the largest full time staff occurs when no contingent workers are used we see that increased flexibility of the contingent workers per se does not reduce full time employment. This phenomenon is surprising given that it is common to view the increased use of contingent workers as a threat to full time employment. Many conclude that anything that would make the contingent workers more attractive, like increased flexibility, could only make the negative impact on full time employment stronger. In Fig. 5 we see evidence that this is not true in general. The reason behind this is that increasing G has two effects. The first effect is that it reduces the number of contingent workers you hire, which may increase N. The second effect is that you want to use the contingent workers

[2] We believe that the non-monotonicity displayed in all the graphs is due to integrality effects.

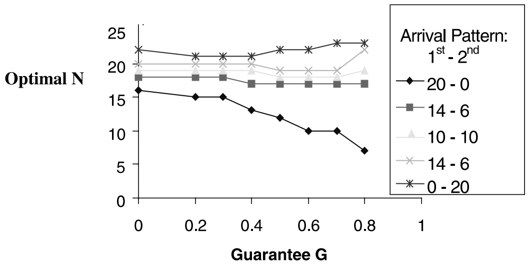

Fig. 5. Regular worker staffing as function of guarantee for different information scenarios.

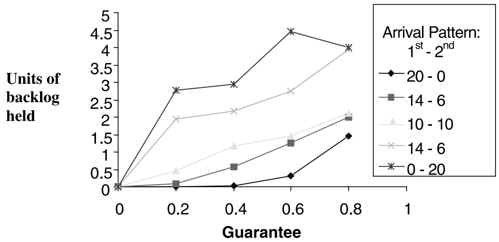

Fig. 6. Units of backlog carried as a function of guarantee for differing information scenarios.

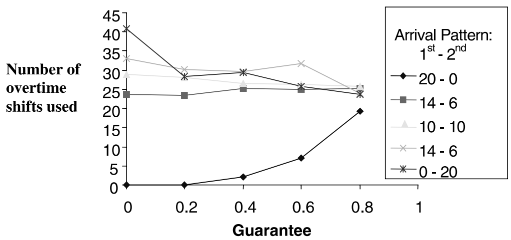

Fig. 7. Number of overtime shifts used as a function of guarantee for differing information scenarios.

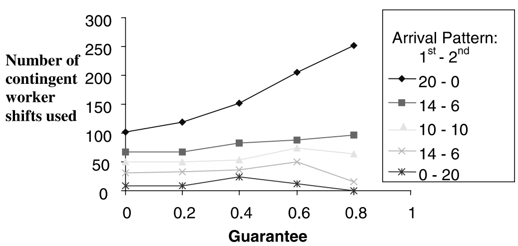

Fig. 8. Number of contingent worker shifts used as a function of guarantee for differing information scenarios.

you have hired more since they have been guaranteed work and this may reduce N. When there is a lot of demand information, e.g., the 20–0 scenario, the second effect is stronger than the first.

In Figs. 6–8 we plot the expected total backlog experienced by the system, the expected total number of overtime shifts used, and the expected total number of contingent worker shifts used, respectively. We see

that backlog tends to increase with less information and less flexibility. This means that both flexibility and information make it easier to handle demand variability and the result is lower costs as is seen in Fig. 2. Surprisingly, we see in Figs. 7 and 8 that overtime and contingent worker usage are in one sense substitutes and in the other complements. When information increases, i.e., moving from scenario 0–20 to scenario 20–0, the number of overtime shifts used declines while the number of contingent workers used increases. However, as flexibility decreases and the guarantee, G, increases, both overtime and contingent worker usage seem to move together. This phenomenon is a result of the way staffing responds to changes in flexibility as was seen in Figs. 3–5.

We have presented results for only one set of cost and productivity parameters. We have conducted experiments in which we vary these parameters and generated predictable results. Increasing the cost or decreasing the productivity of a worker type reduces the staffing of that type and increases costs. Lower backlog penalties lead to higher backlog levels.

The above example allows us to consider the computing and data requirements for an effective implementation of the model presented in this paper. Overall we do not find the model to have data needs beyond those managers typically deal with. In theory, the evaluation of different staffing arrangements with the model can be computationally intensive since it relies on a dynamic programming model. However, by taking advantage of the results discussed in Section 3 it is possible to solve realistic versions of Problem (P) with a desktop PC in a practical amount of time. For example determining the optimal staffing level for a single information scenario and single value of the flexibility parameter took 70 seconds on a Pentium III with a 650 MHz CPU. Experimenting with the performance for larger-scale problems we found that every doubling of the system size in terms of expected demand and staff sizes led to an approximately 16-fold increase in computation time. The relationship between running time and the parameters V and G is approximately linear. Selecting optimal staffing levels would be a planning problem that is typically solved infrequently and therefore running times of several hours may be acceptable. On a tactical level computing the optimal contingent worker and overtime usage each period for a given staffing level takes only of the order of minutes even for problems several time larger than the examples we have presented.

The parameters of the contingent worker contract: $G, \mathrm{OT}_{\mathrm{cw}}$ and the regular worker parameter $\mathrm{OT}_{\mathrm{rw}}$ are set by the decision maker. The labor cost parameters c_{ot}, c_{cw}, c_{rw}, and c_{f} are also set by the decision maker or directly observed. The productivity rates π_{cw}, π_{ot} although not always accurately determined are quantities most managers are comfortable with estimating. The most difficult parameter to determine is the penalty for backlog c_{b}. This parameter can have two components to it, a concrete operational cost and a less-defined cost as perceived by the firm for providing poor service. In practice we believe that a manager should ask herself how much the firm should pay to prevent a unit of backlog with certainty, i.e., eliminate the service delay of a group of customers. This assessment should be a reasonable approximation of the penalty for backlog. Since it is straightforward to use the model to calculate expected backlog c_{b} can also be viewed as a parameter that is adjusted until the system operates with an acceptable level of backlog. The final period backlog penalty, c_{B}, can be calculated in a similar manner taking into account any special end of planning horizon concerns about backlog. The demand process can be estimated using historical data, as can estimates of the absenteeism probability $(1 - p_N)$. Any changes in the notification times for contingent workers would typically require a new analysis of the demand process unless the demand process were a stationary Poisson process.

5. Conclusions

The management of human resources is one of the most complicated managerial tasks and it is impossible to include all issues relevant to staffing decisions in a single model. We have not modeled turnover, training costs or labor supply all of which affect the long term success of any staffing strategy. We have

however isolated a primary motivation for utilizing contingent labor, namely demand uncertainty, and developed a model that can quantify many of the operational costs of utilizing contingent labor under a wide range of circumstances. By using our model a manager can determine the expected short-term cost benefits of working time flexibility derived by better matching labor supply to demand. Broader organizational issues related to contingent labor can be considered in light of these cost benefits.

We have seen that including the timing of demand information in the model yields a better understanding of the value and impact of labor flexibility. In the numerical experiments we found that information and flexibility act as complements. This means that increasing labor flexibility will not yield better performance unless the appropriate information systems are in place to gather sufficient demand information.

One way to interpret the information scenarios is as different amounts of warning time given to contingent workers about their work hours for the next period. In the minimal information case workers are given the most warning time. Full information gives workers the least advance warning. The numerical examples show how the model can be used to quantify the value of work load information or equivalently the value of increased flexibility of workers with regard to notification of work hours. As a result, the model can potentially serve as a useful tool to resolve, in an equitable manner, the inherent tensions in employers and employees needs vis-à-vis working time flexibility.

Finally we have also shown that the model can be used to solve non-trivial problems with reasonable data and computation requirements. This tractability enables the manager to assess the performance of a variety of potential contingent labor contracts efficiently and to make quick decisions about their use from period to period.

Appendix A. Proofs

To simplify the exposition we list a few elementary results on convexity that we make use of in the following proofs:

Property A.1 (Theorem 5.3 in Rockafellar, 1970). *If $f(x,u)$ is jointly convex in (x,u) and (x,u) are in a convex set S, then $\min_u f(x,u)$ is convex in u.*

Property A.2 (Theorem 5.1 in Rockafellar, 1970). *If $f(x)$ is a non-decreasing convex function and $g(x)$ is a convex function, then $f(g(x))$ is a convex function of x.*

Property A.3 (Theorem 5.5 in Rockafellar, 1970). *The positive part ($[\cdot]^+$) operation preserves convexity.*

Property A.4. *If $f(x,y)$ is convex in x and y, then the expectation of $f(x,y)$ over y is convex in x. This fact follows from the fact that multiplication by a positive scalar and summation preserve convexity.*

Property A.5 (Appendix A.5 in Bertsekas, 1976). *If $f(x)$ is convex, then the expectation of $f(x+b)$ over b is convex in x.*

Lemma A.1. *$f_{2V}(x_{2V},\kappa_{2V})$ is a convex function of (x_{2V},κ_{2V}).*

Proof. The final stage, $2V$ is an even stage in which we make an overtime decision. When $c_{\mathrm{B}} \leqslant c_{\mathrm{ot}}/\pi_{\mathrm{ot}}$, we never utilize overtime

$$\Rightarrow \quad f_{2V}(x_{2V},\kappa_{2V}) = c_{\mathrm{B}}x_{2V} \text{ a convex function.}$$

When $c_{\rm ot}/\pi_{\rm ot} \leqslant c_{\rm B}$, we use as much overtime as we can to prevent any backlog

$$\Rightarrow \quad f_{2V}(x_{2V}, \kappa_{2V}) = c_{\rm ot}{\rm Min}\left[{\rm OT}(S), \frac{x_{2V}}{\pi_{\rm ot}}\right] + c_{\rm B}[x_{2V} - {\rm OT}(S)\pi_{\rm ot}]^{+}.$$

For $x_{2V} \leqslant \pi_{\rm ot}{\rm OT}(S)$ we have that $f_{2V}(x_{2V}, \kappa_{2V})$ is increasing at rate $c_{\rm ot}/\pi_{\rm ot}$. For $x_{2V} > \pi_{\rm ot}{\rm OT}(S)$ it is increasing at rate $c_{\rm B}$. Since $c_{\rm B} > c_{\rm ot}/\pi_{\rm ot}$ by assumption, f_{2V} is convex. Note also that $f_{2V}(x_{2V}, \kappa_{2V})$ is independent of κ_{2V}. $\square$

Lemma A.2. *$f_{2V-1}(x_{2V-1}, \kappa_{2V-1})$ is a convex function of $(x_{2V-1}, \kappa_{2V-1})$.*

Proof. In stage $t = 2V - 1$, an odd stage, we make contingent worker use decisions

$$\Rightarrow \quad f_t(x_t, \kappa_t) = \min_{u_t}\{c_{\rm cw}[u_t - \kappa_t]^{+} + E_\theta[f_{2V}([x_t + \theta - u_t\pi_{\rm cw}]^{+}, \kappa_{t+1})]\}.$$

We can rewrite $f_t(x_t, \kappa_t)$ as $\min_{u_t \in [0,M]} H(x_t, \kappa_t, u_t)$.The convexity of $H()$ implies the convexity of $f_t()$, by Property A.1. To complete the proof of the lemma we need to prove that $H()$ is convex in x_t, κ_t, u_t. We can write

$$H(x_t, \kappa_t, u_t) = c_{\rm cw}[u_t - \kappa_t]^{+} + E_\theta[f_{2V}(h(x_t, \theta, u_t), g(\kappa_t, u_t))].$$

By Property A.3 both $h()$ and $g()$ are convex in all their arguments. We have seen that f_{2V} does not depend on κ so we can ignore this dimension. f_{2V} is a non-decreasing function in x since more work cannot reduce costs and by Lemma A.1 it is convex. Therefore, by Property A.2 $f_{2V}(h(x_t, \theta, u_t))$ is convex in (x_t, θ, u_t). Taking expected values with respect to θ also preserves the convexity of $f_{2V}()$ by Property A.4. Since $c_{\rm cw}[u_t - \kappa_t]^{+}$ is convex, $H()$ is convex in (x_t, κ_t, u_t). $\square$

Lemma A.3. *$f_{2V-2}(x_{2V-2}, \kappa_{2V-2})$ is a convex function of $(x_{2V-2}, \kappa_{2V-2})$.*

Proof. When $t = 2V - 2$,

$$f_t(x_t, \kappa_t) = \min_{\omega_t\pi_{\rm ot} \leqslant x_t}\{c_{\rm ot}\omega_t + c_{\rm b}(x_t - \omega_t\pi_{\rm ot}) + E_{\phi'_{t+1}}[f_{t+1}(x_t - \omega_t\pi_{\rm ot} + \phi'_{t+1}, \kappa_t)]\}.$$

Dropping the stage subscript we can write this cost-to-go function as: $f(x, \kappa) = L(x, \kappa) + (c_{\rm ot}/\pi_{\rm ot})x$, where

$$L(x, \kappa) = \min_{[x - {\rm OT}(S)\pi_{\rm ot}]+ \leqslant b \leqslant x}\{\Gamma_t(b, \kappa)\},$$

$$\Gamma_t(b, \kappa) = \left(c_{\rm b} - \frac{c_{\rm ot}}{\pi_{\rm ot}}\right)b + E_{\phi'_{t+1}}[f_{t+1}(b + \phi'_{t+1}, \kappa)] \quad \text{with } b = x - \omega_t\pi_{\rm ot}.$$

By Lemma A.2 we know that $f_{t+1}(x_{t+1}, \kappa_{t+1})$ is convex for $t = 2V - 2$

$$\Rightarrow \quad E_{\phi'_{t+1}}[f_{t+1}(b + \phi'_{t+1}, \kappa)] \text{ is convex in } (b, \kappa) \text{ by Property 5.}$$

Note: $\kappa_{t+1} = \kappa_t$ for even stages t. Since $\Gamma_t(b, \kappa)$ is a sum of convex functions it is convex.

We can now show that $L(x, \kappa)$ is convex as follows.

First, define $x'_0 = \lambda_1 x'_1 + \lambda_2 x'_2$ for $\lambda_1, \lambda_2 \geqslant 0$ and $\lambda_1 + \lambda_2 = 1$ and similarly $\kappa_0 = \lambda_1\kappa_1 + \lambda_2\kappa_2$.

Second, define the following sets: for $i \in [0, 1, 2]$: $S_i \equiv [[x'_i - {\rm OT}(S)\pi_{\rm ot}]^{+}, x'_i]$.

By the definition of $L()$: $L(x'_0, \kappa_0) \leqslant \Gamma_t(b, \kappa_0)$ for all $b \in S_0$ or $L(x'_0, \kappa_0) \leqslant \Gamma_t(\lambda_1 b_1 + \lambda_2 b_2, \lambda_1\kappa_1 + \lambda_2\kappa_2)$ for all pairs (b_1, b_2) such that $b_i \in S_i$.

By the convexity of $\Gamma_t()$:

$$L(x_0', \kappa_0) \leqslant \lambda_1 \Gamma_t(b_1, \kappa_1) + \lambda_2 \Gamma_t(b_2, \kappa_2) \quad \text{for all pairs } (b_1, b_2) \text{ such that } b_i \in S_i$$

$$\Rightarrow \quad L(x_0', \kappa_0) \leqslant \lambda_1 \min_{b_1 \in S_1} \Gamma_t(b_1, \kappa_1) + \lambda_2 \min_{b_2 \in S_2} \Gamma_t(b_2, \kappa_2).$$

Therefore, $L(x, \kappa)$ is convex, and $f_t(x_t, \kappa_t)$ is a sum of convex functions. □

Lemma A.4. *$f_{2V-3}(x_{2V-3}, \kappa_{2V-3})$ is a convex function of $(x_{2V-3}, \kappa_{2V-3})$.*

Proof. In the following we assume that $0 \leqslant u_t \leqslant M$. When $t = 2V - 3$,

$$f_t(x_t, \kappa_t) = \min_{u_t} \{ c_{\text{cw}}[u_t - \kappa_t]^+ + E_{\theta_{t+1}}[f_{t+1}([x_t - u_t \pi_{\text{cw}} + \theta_{t+1}]^+, [\kappa_t - u_t]^+)] \}.$$

We can write $f_t(x_t, \kappa_t)$ as follows (dropping the subscripts): $f_t(x, \kappa) = \min_u \{ \Lambda_t(u, \kappa, x) \}$, where

$$\Lambda_t(u, \kappa, x) = c_{\text{cw}}(u - \kappa) + E_\theta[f_{t+1}([x - u\pi_{\text{cw}} + \theta]^+, 0)] \quad \text{if } \kappa < u$$

and

$$\Lambda_t(u, \kappa, x) = E_\theta[f_{t+1}([x - u\pi_{\text{cw}} + \theta]^+, \kappa - u)] \quad \text{if } \kappa \geqslant u.$$

We define $R1$ to be the set of points in the (u, κ, x) space such that $\kappa < u$, and $R2$ to be the set of points in the same space such that $\kappa \geqslant u$. It is easy to show that within each set, $R1$ or $R2$, $\Lambda_t(u, \kappa, x)$ is convex. We prove that convexity holds across sets.

Assume we have two points $p_1 \in R1$ and $p_2 \in R2$. For non-negative λ_1 and λ_2 such that $\lambda_1 + \lambda_2 = 1$ we define $p_3 = \lambda_1 p_1 + \lambda_2 p_2$ such that $p_3 \in R1$. We first prove that $\Lambda_t(p_3) \leqslant \lambda_1 \Lambda_t(p_1) + \lambda_2 \Lambda_t(p_2)$.

If we define $Q_3 = \Lambda_t(p_3) - \lambda_1 \Lambda_t(p_1) - \lambda_2 \Lambda_t(p_2)$, we have that

$$\begin{aligned} Q_3 = &- c_{\text{cw}}(\lambda_2 \kappa_2 - \lambda_2 u_2) + E_\theta[f_{t+1}([\lambda_1 x_1 + \lambda_2 x_2 - (\lambda_1 u_1 + \lambda_2 u_2)\pi_{\text{cw}} + \theta]^+, 0) \\ &- \lambda_1 f_{t+1}([x_1 - u_1 \pi_{\text{cw}} + \theta]^+, 0) - \lambda_2 f_{t+1}([x_2 - u_2 \pi_{\text{cw}} + \theta]^+, \kappa_2 - u_2)]. \end{aligned}$$

Q_3 can be rewritten as

$$\begin{aligned} Q_3 = &- c_{\text{cw}}(\lambda_2 \kappa_2 - \lambda_2 u_2) + E_\theta[f_{t+1}([\lambda_1 x_1 + \lambda_2 x_2 - (\lambda_1 u_1 + \lambda_2 u_2)\pi_{\text{cw}} + \theta]^+, 0) \\ &- \lambda_1 f_{t+1}([x_1 - u_1 \pi_{\text{cw}} + \theta]^+, 0) - \lambda_2 f_{t+1}([x_2 - u_2 \pi_{\text{cw}} + \theta]^+, 0) + \lambda_2 \Delta_\kappa], \end{aligned}$$

where

$$\Delta_\kappa = f_{t+1}([x_2 - u_2 \pi_{\text{cw}} + \theta]^+, 0) - f_{t+1}([x_2 - u_2 \pi_{\text{cw}} + \theta]^+, \kappa_2 - u_2).$$

Since the most that a unit increase in κ can reduce cost is c_{cw} we can see that $\lambda_2 \Delta_\kappa \leqslant c_{\text{cw}}(\lambda_2 \kappa_2 - \lambda_2 u_2)$.

Because f_{2V-2} is convex and an increasing function of x (more work in system increases costs) we can apply Property A.2 to show that $f_{2V-2}([x - u\pi_{\text{cw}} + \theta]^+, 0)$ is convex in (x, u). Since $\lambda_2 \Delta_\kappa \leqslant c_{\text{cw}}(\lambda_2 \kappa_2 - \lambda_2 u_2)$ we then have that $Q_3 \leqslant 0$.

We now define $p_4 = \lambda_1 p_1 + \lambda_2 p_2$ such that $p_4 \in R2$, and prove that $\Lambda_t(p_4) \leqslant \lambda_1 \Lambda_t(p_1) + \lambda_2 \Lambda_t(p_2)$. If we define $Q_4 = \Lambda_t(p_4) - \lambda_1 \Lambda_t(p_1) - \lambda_2 \Lambda_t(p_2)$ we have that

$$\begin{aligned} Q_4 = &-c_{\text{cw}} \lambda_1 (u_1 - \kappa_1) + E_\theta[f_{t+1}([\lambda_1 x_1 + \lambda_2 x_2 - (\lambda_1 u_1 + \lambda_2 u_2)\pi_{\text{cw}} + \theta]^+, (\lambda_1 \kappa_1 + \lambda_2 \kappa_2 - \lambda_1 u_1 - \lambda_2 u_2)) \\ &- \lambda_1 f([x_1 - u_1 \pi_{\text{cw}} + \theta]^+, 0) - \lambda_2 f_{t+1}([x_2 - u_2 \pi_{\text{cw}} + \theta]^+, \kappa_2 - u_2)]. \end{aligned}$$

For similar reasons as Q_3, we have that $Q_4 \leqslant 0$. This proves that $\Lambda_t(u, \kappa, x)$ is convex. This implies that $\Lambda_t(u, \kappa, x)$ is jointly convex in (u, κ, x) and therefore we can write:

$f_t(x_t, \kappa_t) = \min_{u_t} H(x_t, \kappa_t, u_t)$, where $H(u_t, x_t, \kappa_t)$ is jointly convex in (u_t, x_t, κ_t). The convexity of $H()$ implies the convexity of $f_t()$, by Property A.1. □

Proof of Proposition 1. The proofs of Lemmas A.3 and A.4 only relied on the convexity of the cost-to-go functions in the following periods. These were established in Lemmas A.1 and A.2. Therefore we can apply the above arguments inductively to show that: $f_t(x_t, \kappa_t)$ is convex, for all t. Note in each of the Lemmas A.1–A.4 we have also shown that $f_t(x_t, \kappa_t)$ is the minimum of a convex function of u_t for odd t and ω_t for even t. □

Proof of Proposition 2. $\beta_t(\kappa_t)$ is defined as the backlog b_t that minimizes the function $\Gamma_t(b_t, \kappa_t)$ defined as follows:

$$\Gamma_t(b_t, \kappa_t) = \left(c_{\mathrm{b}} - \frac{c_{\mathrm{ot}}}{\pi_{\mathrm{ot}}}\right) b_t + E_{\phi'_{t+1}}[f_{t+1}(b_t + \phi'_{t+1}, \kappa_{t+1})],$$

where $b_t = x_t - \pi_{\mathrm{ot}}\omega_t$.

Assume that $\beta_t(0) > \beta_t(\kappa_t)$. By definition $\Gamma_t(\beta_t(0), \kappa_t) - \Gamma_t(\beta(\kappa_t), \kappa_t) \geqslant 0$. Since $f_{t+1}(x_t, \kappa_t)$ is increasing in x_t (more work implies higher costs) we have that

$$\left(c_{\mathrm{b}} - \frac{c_{\mathrm{ot}}}{\pi_{\mathrm{ot}}}\right)(\beta_t(\kappa_t) - \beta_t(0)) \leqslant 0.$$

We have assumed that $c_{\mathrm{b}} < c_{\mathrm{ot}}/\pi_{\mathrm{ot}}$, which implies that

$\beta_t(\kappa_t) - \beta_t(0) \geqslant 0$ a contradiction.

The bound on ω_t^* follows from Lemma 1. □

Proof of Proposition 3. The result of Proposition 1 implies that when $\kappa_t = 0$ choosing the optimal u_t in odd stages t, i.e., minimizing $f_t(x_t, 0, u_t)$ for $u_t \in [0, M]$, is a convex minimization problem, where

$$\begin{aligned} f_t(x_t, 0, u_t) &= c_{\mathrm{cw}} u_t + E_{\theta_{t+1}}[f_{t+1}([x_t - u_t\pi_{\mathrm{cw}} + \theta_{t+1}]^+, 0)] \\ &= c_{\mathrm{cw}}\left(u_t - \frac{x_t}{\pi_{\mathrm{cw}}}\right) + E_{\theta_{t+1}}[f_{t+1}([x_t - u_t\pi_{\mathrm{cw}} + \theta_{t+1}]^+, 0)] + c_{\mathrm{cw}}\frac{x_t}{\pi_{\mathrm{cw}}} \\ \Rightarrow \quad \min_{u_t} f_t(x_t, 0, u_t) &= \min_{X_t}\left\{\frac{-c_{\mathrm{cw}}}{\pi_{\mathrm{cw}}} X_t + E_{\theta_{t+1}} f_{t+1}([X_t + \theta_{t+1}]^+, 0)\right\} + c_{\mathrm{cw}}\frac{x_t}{\pi_{\mathrm{cw}}}, \end{aligned}$$

where $X_t = x_t - u_t\pi_{\mathrm{cw}}$ and $X_t \in [x_t - M\pi_{\mathrm{cw}}, x_t]$. The statement of the proposition follows. □

References

Abraham, K., 1986. Flexible staffing arrangements. In: Hart, R.A. (Ed.), Employment, Unemployment and Labor Utilization. Unwin Hyman Inc., USA, pp. 288–312.

American Staffing Association, 2001. Quarterly Staffing Survey 2001. American Staffing Association, Alexandria, VA. Available from www.natss.org.

Anupindi, R., Bassok, Y., 1997. Analysis of supply contracts with total minimum commitment. IIE Transactions 30 (8).

Anupindi, R., Bassok, Y., 1998. Analysis of supply contracts with commitments and flexibility. Mimeo.

Applebaum, E., 1987. Restructing work. In: Hartmann, H. (Ed.), Computer Chips and Paper Clips, Technology and Women's Employment. National Academy Press, Washington, DC.

Berman, O., Larson, R.C., 1994. Determining optimal pool size of a temporary call-in work force. European Journal of Operational Research 1 (73), 55–64.

Bertsekas, D., 1976. Dynamic Programming and Stochastic Control. Academic Press, New York.

Brotherton, P., 1995. Staff to suit. HR Magazine 40 (12), 50–53.
Chen, F., Fedegruen, A., Zheng, Y., 2001. Coordination mechanisms for a distribution system with one supplier and multiple retailers. Management Science 47 (5), 693–708.
Christensen, K., 1995. Contingent Work Arrangements in Family-sensitive Corporations. Center on Work and Family, Boston University, Boston.
Clinton, A., 1997. Flexible labor: Restructing the American work force. Monthly Labor Review 120 (8), 3–17.
Donohue, K.L., 2000. Efficient supply contracts for fashion goods. Management Science 46 (11), 1397–1411.
Manpower Inc., 2000. Annual Report. Manpower Inc., Milwaukee, WI.
Milner, J., Pinker, E., 2001. Contingent labor contracting under demand and supply uncertainty. Management Science 47 (8), 1046–1062.
Nollen, S., Axel, H., 1995. Managing Contingent Workers. American Management Association, New York.
Pinker, E., 1996. Models of flexible workforce management in uncertain environments. Ph.D. thesis, MIT, Cambridge, MA.
Polivka, A.E., 1989. On the definition of 'contingent work'. Monthly Labor Review 112 (12), 9–16.
Rockafellar, R.T., 1970. Convex Analysis. Princeton University Press, Princeton, NJ.
US Bureau of Labor Statistics, 1995. Current Population Survey, February 1995.
Wild, B., Schneeweiss, C., 1993. Manpower capacity planning – A hierarchical approach. International Journal of Production Economics 30/31, 95–106.